

Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function

Ingo Ciolkowski · Dierk Wanke · Rainer P. Birkenbihl · Imre E. Somssich

Received: 3 December 2007 / Accepted: 21 May 2008 / Published online: 4 June 2008
© The Author(s) 2008

Abstract WRKY transcription factors have been shown to play a major role in regulating, both positively and negatively, the plant defense transcriptome. Nearly all studied WRKY factors appear to have a stereotypic binding preference to one DNA element termed the W-box. How specificity for certain promoters is accomplished therefore remains completely unknown. In this study, we tested five distinct Arabidopsis WRKY transcription factor subfamily members for their DNA binding selectivity towards variants of the W-box embedded in neighboring DNA sequences. These studies revealed for the first time differences in their binding site preferences, which are partly dependent on additional adjacent DNA sequences outside of the TTGACY-core motif. A consensus WRKY binding site derived from these studies was used for in silico analysis to identify potential target genes within the Arabidopsis genome. Furthermore, we show that even subtle

amino acid substitutions within the DNA binding region of AtWRKY11 strongly impinge on its binding activity. Additionally, all five factors were found localized exclusively to the plant cell nucleus and to be capable of transactivating expression of a reporter gene construct in vivo.

Keywords Arabidopsis promoters · Nuclear localization · Transactivation · W-box element

Introduction

Transcription factors act in concert with other components of the transcriptional machinery to modulate the expression of target genes in a temporal and spatial manner. In general, they do so by binding to short defined nucleotide motifs (*cis*-elements) within the regulatory regions of genes that are under their control. Different classes of transcription factors have characteristic DNA-binding domains that discriminate between distinct *cis*-regulatory elements at target sites within the genome. Our studies have been focused on a family of zinc-finger type transcription factors, designated WRKY. WRKY factors comprise a large family of DNA-binding proteins found in all plants (Eulgem and Somssich 2007). Although not completely restricted to the plant kingdom, this family has expanded enormously in higher plants whereas they appear to have been lost in yeast and animal lineages (Ülker and Somssich 2004). WRKY proteins have been implicated in the regulation of developmental processes such as trichome and seed development and leaf senescence (Hinderhofer and Zentgraf 2001; Johnson et al. 2002; Luo et al. 2005), but their major functions appear to be in helping the plant to cope with various abiotic and biotic stresses (Ülker and Somssich 2004; Journot-Catalino et al. 2006; Li et al.

Electronic supplementary material The online version of this article (doi:10.1007/s11103-008-9353-1) contains supplementary material, which is available to authorized users.

I. Ciolkowski · D. Wanke · R. P. Birkenbihl ·
I. E. Somssich (✉)
Department of Plant Microbe Interactions, Max Planck Institute
for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829
Köln, Germany
e-mail: somssich@mpiz-koeln.mpg.de

Present Address:

I. Ciolkowski
Justus-Liebig-Universität Giessen, IPAZ, Heinrich-Buff-Ring
26-32, 35392 Giessen, Germany

Present Address:

D. Wanke
ZMBP – Pflanzenphysiologie, Auf der Morgenstelle 1, 72076
Tübingen, Germany

2006; Wang et al. 2006; Xu et al. 2006; Zheng et al. 2006; Shen et al. 2007).

All WRKY proteins contain a 60 amino acid long peptide region, termed the WRKY domain, which constitutes their DNA binding regions. Apart from the invariant name-giving amino acid residues W R K and Y, this domain also contains conserved cysteine and histidine residues that bind one zinc atom and form a finger-like structure. Both the WRKY residues as well as the zinc finger motif are required for proper DNA binding of the protein (Maeo et al. 2001). Based on the number of WRKY domains and the pattern of the zinc-finger motif, WRKY factors have been classified into three major groups (Eulgem et al. 2000). The number of family members in higher plants can range from 74 in Arabidopsis, 98 to 102 in rice or even more in other species such as tobacco (Ülker and Somssich 2004; Ross et al. 2007).

Gel-shift experiments, random binding site selection, DNA-ligand binding screens, yeast one-hybrid studies and cotransfection assays performed with different plant WRKY proteins have illustrated that the *cis*-element 5'-TTGAC-C/T-3', termed the W-box, represents the minimal consensus required for specific DNA binding (de Pater et al. 1996; Rushton et al. 1996; Wang et al. 1998; Chen and Chen 2000; Cormack et al. 2002). Only in the case of the barley WRKY factor SUSIBA2 and the tobacco NtWRKY12 have additional DNA binding sites been identified, although SUSIBA2 also shows strong *in vitro* binding to the W-box motif (Sun et al. 2003; van Verk et al. 2008). Considering the size of the WRKY family within a given species and the apparent stereotypic binding preference of these proteins to the W-box, it is difficult to foresee how specificity for certain promoters is accomplished. A certain level of specificity may be conferred by additional nucleotide sequences flanking W-box elements with respective gene promoters. Furthermore, various WRKY subgroup family members within a single plant species may also differ somewhat in their DNA binding requirements. Both of these aspects have not yet been directly addressed for WRKY transcription factors. Finally, the involvement of WRKY factors in diverse higher-order nucleoprotein complexes can also be assumed and may in fact be the major criteria in determining promoter selectivity and transcriptional output.

In this study, we selected five Arabidopsis WRKY transcription factors representing the three major groups of WRKY proteins including three from the largest, namely group II. These factors were analyzed with respect to their *in vitro* binding capabilities to various W-box variants and promoter sequences containing W-box motifs. This analysis already revealed clear as yet not observed binding site preferences between certain representatives. Furthermore, we show that all studied WRKY factors are localized

exclusively to the nucleus and that all have *in vivo* trans-activation capabilities.

Materials and methods

Recombinant fusion proteins

The AtWRKY cDNA fragments were obtained with specific Gateway-primers in RT-PCR reactions and cloned in-frame into the modified bacterial expression vector pGEX2T (Pharmacia) or into the vector pQE30a (Qiagen) and verified by sequencing. The C-terminal protein construct of WRKY11 (WRKY CTD) consisted of 98 amino acids encompassing the WRKY DNA binding domain and in addition carrying a 6×His tag at its N-terminus MRGSHHHHHMKRTVRVPAISAKIADIPPDEYSWRKYGQKPIKGSPPHPRGYKCSSTFRGCPARKHVERALDDPAMLIIVTYEGEHRHNQSAMQENISSSGINDLVFASA. Mutations within the WRKY11 CTD were introduced using the megaprimer method (Ke and Madison 1997) and cloned into the Gateway compatible vector pASK5 (IBA). A list of the primers used in this study are given in the supplementary table (S-Table 1). Protein expression was carried out in *E. coli* strain BL21(DE3) following addition of IPTG (1 mM) or 2 µg/µl anhydrous tetracyclin (IBA) for 3–4 h. Bacteria were harvested by centrifugation. The pellet was resuspended and lysed under native conditions but including protease inhibitors at 4°C (PIERCE, lysis protocol according to the instructions of the manufacturer). Due to technical difficulties encountered in purifying the respective recombinant WRKY proteins all subsequent experiments were performed using total soluble *E. coli* protein extracts derived from bacteria expressing the WRKY cDNA construct or the control empty vector construct. For western blot analyses proteins were separated by SDS-PAGE and blotted onto a PVDF membrane (Sartorius). The membranes were probed with *anti*-GST antibody (Pharmacia) or anti-6×HIS antibody and anti-goat or anti-mouse IgG conjugated to alkaline phosphatase, respectively. *Anti*-strep antibody was directly conjugated to alkaline phosphatase. Membranes were developed with NBT-BCIP solution Sigma-fast (Sigma).

Electrophoretic mobility shift assay (EMSA)

Equivalent amounts of sense and antisense fragments of the respective oligonucleotides (Invitrogen) were annealed in 40 mM Tris-HCl pH 7.5, 20 mM MgCl₂, and 50 mM NaCl starting at 95°C and allowing to cool slowly to room temperature. About 1.5 µl of the solution corresponding to 150 ng DNA was used for end-labeling with 5 U of terminal

transferase (Roche) and 20 μ Ci dCTP (Amersham) in a total volume of 25 μ l according to the manufacturer's instructions (Roche). The probes were purified on Sephadex G-25 columns and diluted to 5000 cpm/ μ l. Binding reactions were performed according to Maeo et al. (2001). Reaction mix which included 25 μ g of total soluble bacterial protein extract was incubated for 20–25 min to 25°C after addition of the radioactively labeled probe then applied to the gel. Gel electrophoresis was carried out at 4°C. Sequence of 1 \times W2: TTATTCAGCCATCAAAAGTTGACCAATAAT. For competition experiments unlabelled DNA was added as competitor to the binding reaction and incubated for 10 min on ice. Labeled probe DNA was added and the sample was incubated at room temperature for another 15 min prior to loading on the gel. For all EMSA experiments described in this paper, the specificity of the observed signals obtained with the labeled probes was also confirmed using an excess (50- to 1000-fold) of unlabeled competitor DNA (data not shown).

Biolistic transformation of leaves by particle bombardment

Arabidopsis thaliana plants (ecotype Columbia-0) were grown under long day conditions (18 h light, 6 h darkness, 20°C). Leaves of 4–5-week-old plants were detached, placed on 0.5 \times MS plates and incubated in a light chamber for at least 4 h prior to bombardment. On average 40 to 50 leaves were used per assay.

The 4 \times W2::GUS reporter contains a tetramer of the W-box motif TTGACC (Eulgem et al. 1999). The *SIRKp*::GUS construct was described by Robatzek and Somssich (2002). Full-length cDNAs of WRKY6, WRKY11, WRKY26, WRKY38 and WRKY43 were fused in-frame to the GFP reporter gene and expressed under the control the CaMV35S promoter (p35S-GW-GFP). For transient expression in onion cells, 5 μ g of the respective constructs were introduced into epidermal cells via particle bombardment. Subcellular localization was microscopically monitored 20–24 h post bombardment. The expression constructs were introduced into *Arabidopsis* leaves via particle bombardment (PDS-1000/He Biolistic® Particle Delivery System, Bio-Rad), essentially as described by Robatzek and Somssich (2002). Briefly, 50 μ l of gold particles (1 μ m diameter) were coated with both the reporter and effector plasmids (3 μ g DNA each). When using one reporter plasmid and two different effector plasmids, 2 μ g of each DNA was used. About 3 mg of gold particles were delivered per shot. After 24 h bombardments, the leaves were infiltrated with GUS staining solution for 3–5 min and incubated at 37°C overnight followed by clearing of the leaves with 95% ethanol. Reporter gene activity staining was evaluated macroscopically, taking into account all leaves used in the experiment.

Cis-element analysis

The GenBank DNA sequence flat-files were downloaded from the Entrez Plant Genomes Central at NCBI (<http://www.ncbi.nlm.nih.gov/>; NC_003070.5, NC_003071.3, NC_003074.4, NC_003075.3, NC_003076.4). Sequences of either 1500 or 600 bp upstream of the annotated translation start sites (ATG) were extracted and compiled to FASTA formatted text with the aGenBankSQL script v3.3.4 (Berendzen et al. 2006), using the default code modifications. Number of W-box motifs has been assessed counting their occurrences in all chromosome sequences and the promoters. The fold-differences were calculated on the number of W-box occurrences in the two promoter datasets divided by the average number of W-boxes in the genome (number of W-boxes divided by the total number of bases of all chromosomes).

W-box elements and derived consensus were examined with the Motif Mapper Open Source Script Package (<http://www.mpiz-koeln.mpg.de/coupland/coupland/mm3/html/>) as has been described previously (Berendzen et al. 2006). The Sequence logos for the two W-box consensus motifs were derived from a position weight matrix using WebLogo (<http://weblogo.berkeley.edu/>; Crooks et al. 2004).

Functional categorization of the gene lists was performed using the Gene Ontology annotation form at TAIR (http://www.arabidopsis.org/portals/genAnnotation/functional_annotation/go.jsp).

To assess values for significance (*P*-values), we retrieved the GO-annotations for all genes using the TAIR7 release of the *Arabidopsis* genome. A set of 20669 GO-annotations was returned, which could be used as a background model to calculate the hypergeometric distribution.

The equation used to calculate the hypergeometric distribution is

$$P(X = x) = \frac{\begin{bmatrix} M \\ x \end{bmatrix} \begin{bmatrix} N - M \\ n - x \end{bmatrix}}{\begin{bmatrix} N \\ n \end{bmatrix}}$$

where *N* is the size of the population with *M* as the number of incidents within the population and *n* is the size of the sample with *x* as the number of incidents within the sample.

Results

DNA binding selectivity of *Arabidopsis* WRKY factors

To date, all studied plant WRKY transcription factors show high binding preference to the DNA sequence element, 5'-C/TTGACT/C-3', termed the W-box (Ülker and Somssich 2004; Eulgem and Somssich 2007 and citations

therein). However, no systematic study has been reported testing whether members of the different WRKY subgroups within a single plant species all show similar DNA binding requirements. Thus, we choose five selected subgroup representatives of the *Arabidopsis thaliana* WRKY gene family for qualitative DNA binding studies. These members represent the WRKY groups I (WRKY26; At5g07100), IIb (WRKY6; At1g62300), IIc (WRKY43; At2g46130), IId (WRKY11; At4g31550) and III (WRKY38; At5g22570) (Eulgem et al. 2000). Full-length cDNAs for all representatives served as templates for expression in *E. coli* and total soluble protein extracts derived from these bacteria were directly used for further analysis. Initially, we encountered several problems attempting to express a large set of WRKY proteins (both as GST fusions or epitope tagged variants) in *E. coli*. Often, their expression proved detrimental for bacterial growth. Since WRKY factors are zinc-finger proteins, their expression may in some way negatively influence zinc homeostasis. Exogenous addition of zinc however did not relieve this problem. In several instances bacterial growth was not significantly influenced but the specific WRKY proteins were found exclusively in inclusion bodies. No attempts were made to purify the proteins from this source since previous work in our laboratory has shown that during the denaturation and subsequent renaturation steps misfolding of these proteins often occurs resulting on loss of W-box binding ability. Affinity purification of epitope-tagged WRKY proteins under mild condition also proved problematic. Thus, only those five WRKY members were taken into consideration for further analyses for which sufficient protein was present in the total soluble bacterial protein extracts. As controls for all experiments, protein extracts derived from identically treated bacteria harboring only the appropriate empty expression vector cassette were always included. In no case did these control extracts result in detectable sequence-specific binding to the tested DNA sequences (data not shown).

In the first set of experiments, we synthesized an oligonucleotide (1×W2; 5'-TTATTCAGCCATCAAAGTTGACCAATAAT-3') identical in sequence to the region of the parsley *PR1-1* promoter containing one W-box (W2; Rushton et al. 1996). The W2 element was previously shown to mediate strong elicitor-dependent gene activation and to be bound specifically by WRKY factors (Rushton et al. 1996, 2002). In addition, several variants of this oligonucleotide were synthesized containing single or double nucleotide base exchanges within or directly adjacent to the W2 box (Fig. 1a). The various double stranded oligomers were radioactively labeled, incubated with the different bacterial protein extracts, and specific DNA binding by the respective WRKY factors detected by electrophoretic mobility shift assays (EMSA; Fig. 1b).

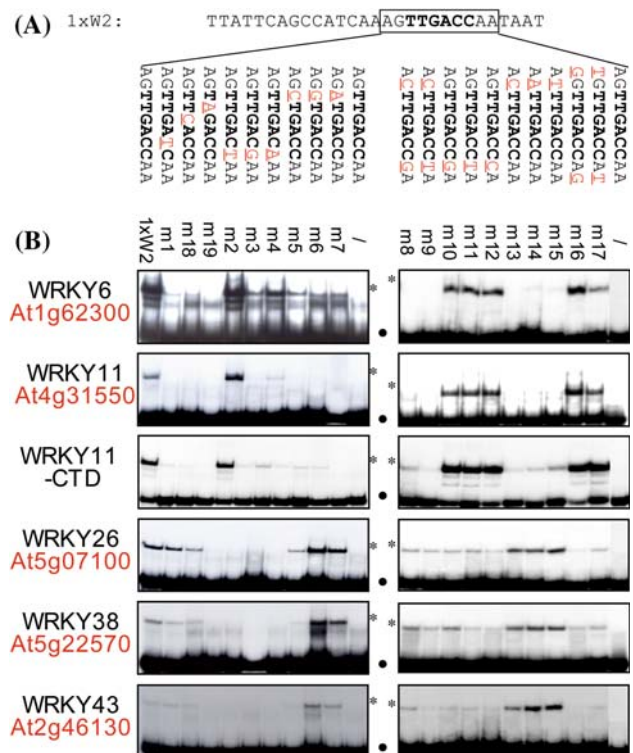


Fig. 1 Binding site preferences of different Arabidopsis WRKY subgroup representatives to W-box variants. (a) Sequence of the parsley *PR1-1* promoter region designated 1×W2 (Rushton et al. 1996). The W-box element is in bold letters. Box marks the region of the oligonucleotide within which base substitutions were generated to produce the variants listed below. The W2 box element is highlighted in bold whereas the base alterations are plain letters and underlined. (b) EMSA of the radioactively labeled designated W2 box variants incubated with crude bacterial extracts containing the indicated recombinant WRKY protein. Lanes marked by a slash contain only the labeled free DNA probes with no protein added. Specific retarded DNA-protein complexes are marked by open asterisks in the middle of the composite, whereas dots designate positions of the free probes

Surprisingly, two discrete binding site preferences could be observed with the five WRKY representatives used (Fig. 1b). In the case of WRKY6 and WRKY11, clear binding was detected with the two known functional W-box variants TTGACC (1×W2) and TTGACT (m2), whereas no specific binding was seen with oligomers containing mutations within this W-box consensus (m1, m18, m19, m3 to m7). The binding of these two factors was similarly influenced negatively by mutations of the guanosine residue directly 5' adjacent of the W-box motif (m8, m9, m14 and m15). For both factors, base exchanges 3' of the element did not seem to affect binding of both factors (m10 to m12, m16, m17). However, in the case of m16 and m17 we cannot completely rule out that simultaneous exchanges at the 5' ends may have compensated for the base exchanges at the 3' ends of these elements. Identical binding properties were observed using the entire

WRKY11 protein or only the 98 amino acid long region harboring the WRKY domain that mediates DNA binding (Fig. 1b, WRKY11-CTD). In contrast, WRKY26, WRKY38 and WRKY43 showed common binding characteristics that clearly differed from those observed for WRKY6 and WRKY11. All three proteins bound the TTGACC element (1×W2) but showed no binding to the other functional W-box variant TTGACT (m2). Exchange of the first 5' base within the element from T to G or A actually enhanced binding (Fig. 1b; m6, m7). Lower but clearly detectable binding was still observed with other single base substitutions (m1, m18 and m5). Importantly, mutations of the G residue directly 5' adjacent of the W-box element had a clear positive effect on binding (Fig. 1b; m13 to m15). Since the levels of bacterially expressed individual WRKY factors varied significantly as determined by western blots (data not shown), binding strength comparisons between the factors were not made. Rather, only qualitative differences were scored. Still, we can conclude that the five tested WRKY proteins fall into two distinct groups (WRKY6 and 11 on the one hand, and WRKY26, 38 and 43 on the other) showing different DNA sequence requirements for binding. Furthermore, the presence of a W-box element in a DNA region alone is not sufficient for WRKY factor binding, but can be influenced by additional sequences in its neighborhood. By 5' and 3' deletions of the original 1×W2 oligomer we determined that 4 bp proximal to and 3 bp distal of the W-box motif are required for efficient WRKY factor binding, but that the orientation of the element within the sequence does not influence binding (data not shown).

WRKY factor binding to diverse W-box containing Arabidopsis promoters

We next tested whether W-box containing sequence regions of native Arabidopsis gene promoters will also yield additional information concerning selective binding requirements of WRKY factors. For this, four promoters were selected derived from the *SIRK/FRK1* (At2g19190), *CMPG1* (At3g02840), *4CL4* (At3g21230) and *WRKY11* loci. The *SIRK/FRK1* gene encodes a receptor kinase whose expression is up-regulated during leaf senescence and upon pathogen challenge (Asai et al. 2002; Robatzek and Somssich 2002). Its promoter contains numerous W-box motifs and several WRKY factors have been implicated in its regulation. *CMPG1* represents an immediate-early pathogen-responsive gene coding for a U-box protein that is required for efficient activation of defense mechanisms in tobacco and tomato (Gonzalez-Lamothe et al. 2006). The W-box-containing region was shown to mediate this response (Heise et al. 2002). *4CL4* codes for a 4-coumarate:CoA ligase, a key enzyme in general phenylpropanoid metabolism with unusual catalytic properties (Hamberger

and Hahlbrock 2004). The occurrence of three TATA-proximal W-box elements is an absolute singularity among all known *4CL* gene promoters.

WRKY11 and WRKY26 were chosen as representatives of the two distinct binding type factors for further gel shift studies. Oligomers spanning defined W-box containing regions of the selected promoters were synthesized (Fig. 2a) and used as probes. Figure 2b shows the results of these experiments. For the majority of cases, the outcome of the interaction was predictable based on the information gained from Fig. 1b. For example the two W-box elements present in the −731 to −705 region of *AtSIRK* promoter (Fig. 2a, 1) were expected to both be bound by WRKY11 and thus yield two retarded complexes (one site occupied or both sites) in EMSA (Fig. 2b, lane 1). Similar results are observed for the −47 to −31 *AtSIRK* promoter region probe (Fig. 2, lane 4). In contrast, little binding is seen with the two other analyzed regions (Fig. 2, lanes 2 and 3). Both of these regions contain one inverted W-box core and one W-box motif in a sequence environment detrimental for WRKY11 binding (compare to Fig. 1, m4, m13 and m15). WRKY11 binds to all other promoter regions analyzed with the exception of the *At4CL*-W2 (Fig. 2, lane 10). In *At4CL*-W2, the TTGACC sequence is directly preceded by

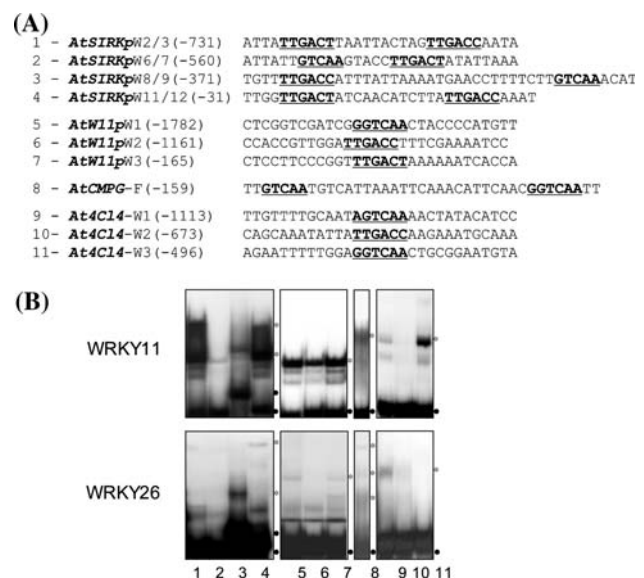


Fig. 2 Binding of WRKY factors to W-box containing promoter sequences. **(a)** Upper strand sequences of the oligonucleotide probes synthesized for gel shift assays. The sequences represent different W-box containing regions of four selected Arabidopsis gene promoters as described in the text. W-box motifs are highlighted in bold. **(b)** EMSA of the indicated radioactively labeled probes incubated with crude bacterial extracts containing WRKY11 or WRKY26 recombinant protein. Lane numbering is identical to the probe numbers in **a**. The specific retarded DNA-protein complexes are marked by open asterisks at the right-hand side of each gel whereas dots designate positions of the free running probes

an A residue shown to have a negative effect on binding (Fig. 1, m14). In the case of WRKY26, clear binding was only observed with the *AtSIRK* promoter probe 3. This was anticipated (see Fig. 1, m15). Formation of the retarded complexes depicted in Fig. 2 could be competed away by an excess of the corresponding non-labeled probes (data not shown). One must note however, that not all observed interactions could be predicted based on our limited studies, indicating that additional neighboring nucleotide requirements are needed for proper WRKY factor binding. Nevertheless, these data again illustrate that WRKY proteins do have selective binding preferences to certain W-box motifs embedded within different promoter sequences.

Binding of Arabidopsis WRKY factors to W-box dimers

WRKY factors have been shown to bind to promoter regions often containing closely spaced W-box motifs (Rushton et al. 1996; Eulgem et al. 1999; Yang et al. 1999; Yu et al. 2001; Chen and Chen 2002; Marè et al. 2004; Zhang et al. 2004).

We therefore investigated how alterations in the spacing of two W-box elements influences binding with the five Arabidopsis WRKY factors. As a starting point we generated a dimer (2×W2s24) of the 1×W2 oligomer, which was bound by all five WRKY factors (Figs. 1b and 3a). 2×W2s24 contains two W-box motifs separated by 24 bp. Four additional derivatives of the dimer were synthesized differing in the number of base pairs between the two W-box elements (Fig. 3a). Gel shift assays revealed that all double-stranded 2×W2s24 derivatives were bound by the WRKY proteins (Fig. 3b). Interestingly however, with the exception of WRKY26, all WRKY factors mainly bind to only one of the two potential W-box sites within the DNA oligomer. Simultaneous binding to both sites, yielding an additional more slower migrating shifted protein-DNA complex in the gel, occurs to only a limited extent except for WRKY26. Using only the WRKY11 DNA binding domain (WRKY11 CTD; Fig. 3b) in such experiments yields stronger signals for the higher molecular weight complex indicating the size restrictions may partly limit simultaneous binding to both sites. This is further supported by the fact that this larger retarded complex is no longer observed when the W-box motifs are separated by only a single base pair (Fig. 3b, 2×W2s1). That the formation of both retarded complexes is dependent on W-box specific binding was demonstrated by competition experiments using either an excess of non-labeled 2×W2s24 or a variant (2×W2m, Fig. 3a) in which both W-box motifs were altered from TTGACC to TacACC. As illustrated for WRKY11-CTD, both complexes are competed away by an excess of the wildtype sequence, whereas this is not the case using 2×W2m (Fig. 3c). Similar

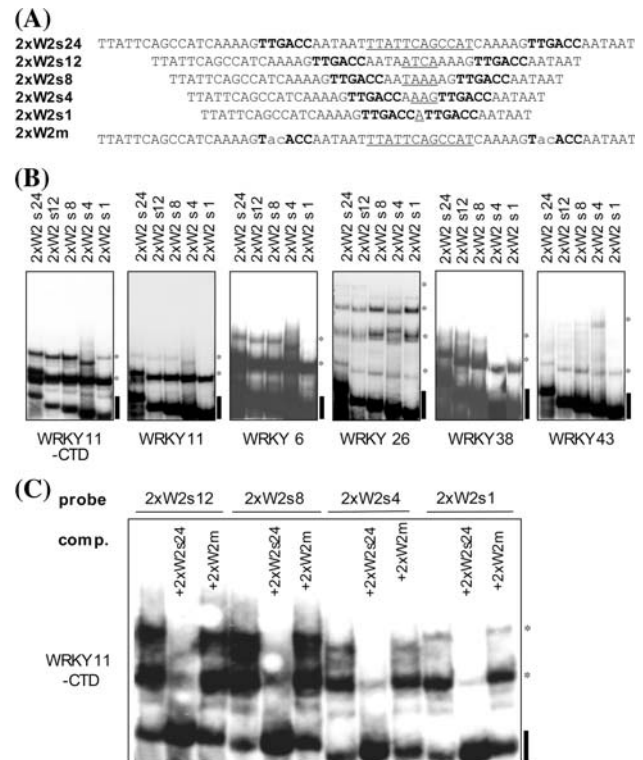


Fig. 3 Spacing between two W-boxes does not significantly influence WRKY binding. **(a)** Upper strand sequences of the oligonucleotide probes synthesized for gel shift assays. **(b)** EMSA of the indicated radioactively labeled probes incubated with crude bacterial extracts containing the respective WRKY recombinant proteins. **(c)** EMSA with WRKY11 CTD protein and radioactively labeled DNA probes as indicated, in the presence or absence of non-labeled competitor DNA. A 500-fold excess of the respective competitor (marked in the lanes) was included in the binding reaction. Specific retarded DNA-protein complexes are marked by open asterisks at the right-hand side of each gel, whereas black bars designate the positions of the free running probes

competition results were obtained for the other WRKY factors (data not shown).

Amino acid residues influencing DNA binding activity of the WRKY domain

Maeo et al. (2001) demonstrated for tobacco WRKY9 that the two cysteine and histidine residues forming the zinc-finger motif and the highly conserved WRKYGQK amino acid stretch within the WRKY domain are important for W-box binding activity. Furthermore, they and others have shown that despite high sequence identity between the two WRKY domains of group I WRKY factors, only the C-terminal WRKY domain (CTD) actually contributes significantly to DNA binding (de Pater et al. 1996; Eulgem et al. 1999; Maeo et al. 2001).

We expressed variants of the CTD domain of Arabidopsis WRKY11 in *E. coli* and used these bacterial extracts

for EMSA to define additional amino acid residues within the WRKY domain important for DNA binding. *At*WRKY11-CTD showed identical binding characteristics towards W-box containing sequences as the full-length WRKY11 protein (compare Figs. 1–4b) but yielded higher and consistent amounts of recombinant protein. Figure 4a shows the peptide sequence of the *At*WRKY11 WRKY domain (*At*WRKY11-CTD) and the various, mostly single amino acid substitutions that were generated within this domain. Exchange of selective amino acid positions was based on a pile up comparison between all Arabidopsis WRKY domains and in particular, considering differences observed between the N-terminal and C-terminal domains of group I members. The bacterially expressed recombinant proteins all carried a 6×His- or Strep-tag for detection and

quantification on western blots, which verified that all proteins were produced (data not shown). As expected, amino acid exchanges within the WRKY motif itself dramatically affected binding to all W-box containing DNA sequences (Fig. 4, m1, m3, m4 and m5) consistent with previous reports (Maeo et al. 2001; Duan et al. 2007). Similarly, a C to G substitution, potentially disrupting the zinc-finger motif (Fig. 4a, m2), also strongly affected DNA binding, but did not completely abolish it (Fig. 4b, m2). Using the *AtSIRKp*W11/12 promoter region containing two W-boxes as labeled probe, all additional WRKY11-CTD substitution variants were tested via EMSA (Fig. 4c). In most cases, single amino acid alterations negatively affected DNA binding. Only in three instances was binding capability unaffected (Fig. 4c, m7, m16 and m17). For the variants m8 and m17, in which the same glycine residue was substituted by either a serine or a histidine, only the G to H exchange, resulting in a substitution of an uncharged polar amino acid to a basic charged one, had an effect. It is important to note however, that for the majority of variants, we choose to exchange amino acids sharing similar properties. For example, in the variants m7, m10, m12, and m18, arginine was substituted by lysine or vice versa (Fig. 4a). Still, such substitutions often profoundly affected DNA binding ability, suggesting that many amino acid residues within the WRKY domain are critical for proper binding function. Interestingly, one substitution (D to E; Fig. 4c, m6) outside of the WRKY domain also strongly influenced W-box binding. This is in accordance with results from tobacco demonstrating that deletion of 4 residues just N-terminal to the WRKY domain of *Nt*WRKY9 completely abolished DNA binding (Maeo et al. 2001) and with data from Babu et al. (2006) revealing a plant-specific zinc cluster directly preceding the *At*WRKY11-CTD.

In vivo interaction of WRKY factors with W-box elements

All five tested WRKY factors showed in vitro binding activity to the 1×W2 box motif independent of the orientation of the element within the DNA sequence. Since individual WRKY factors have been shown to function as activators or repressors of transcription (Robatzek and Somssich 2002; Zhang et al. 2004) but also to lack such activity (Hara et al. 2000), we tested the ability of the Arabidopsis WRKY subgroup members to activate W-box-dependent transcription of a reporter gene *in vivo*. We first fused the *gfp* reporter gene in frame C-terminal to the respective full-length WRKY cDNAs. Expression of these constructs was driven by the strong constitutive 35S CaMV promoter. In transient transfections using epidermal onion cells, detection of GFP fluorescence confirmed that the WRKY-GFP fusion proteins were synthesized and

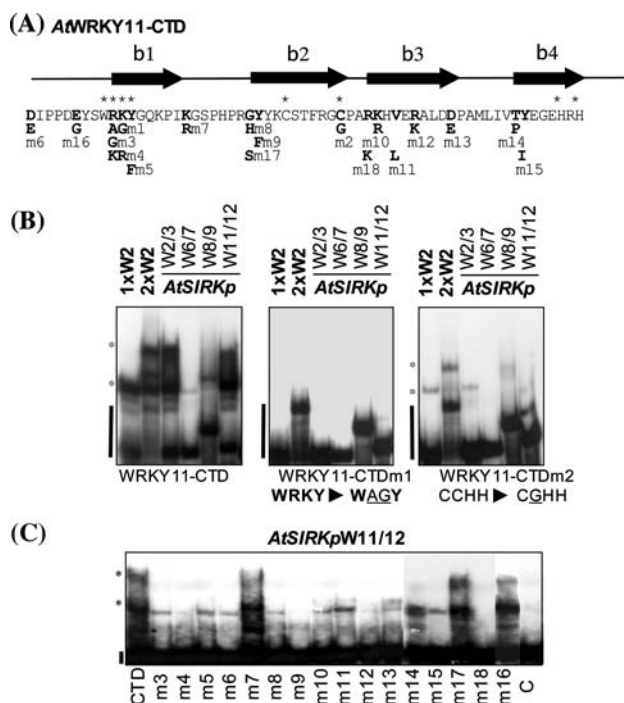


Fig. 4 Defining critical amino acid residues of *At*WRKY11 for DNA binding. (a) Peptide sequence of the C-terminal WRKY domain of WRKY11 (*At*WRKY11-CTD). Highlighted in bold letters are those amino acid residues within the peptide for which substitution variants (m1 to m18) were generated. The exchange amino acids within the individual variants are given below. The open stars above the peptide sequence mark the positions of the invariant WRKY stretch and the cysteines and histidines in the zinc finger motif that are conserved in WRKY transcription factors. Black arrows above the sequence indicate regions of β -strands as recently determined for the *At*WRKY4-CTD (Yamasaki et al. 2005). (b) EMSA of crude bacterial extracts harboring the indicated WRKY11-CTD proteins having amino acid exchanges at conserved positions with different radioactively labeled W-box containing probes. (c) EMSA of crude bacterial extracts harboring the various WRKY11-CTD protein variants given in a, with radioactively labeled *AtSIRKp*W11/12 probe. Specific retarded DNA-protein complexes are marked by open asterisks at the left-hand side of each gel, whereas black bars designate the positions of the free running probes

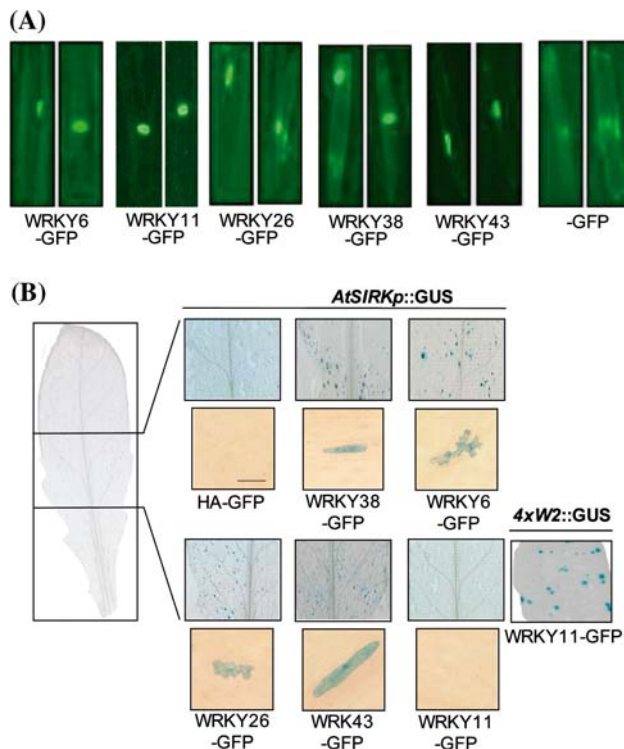


Fig. 5 Nuclear localization and in vivo transactivation functions of different WRKY subfamily representatives. **(a)** Transient expression of WRKY-GFP fusion proteins in onion epidermal cells. Two individual cells transfected with the indicated expression constructs and showing GFP fluorescence are shown. WRKY6-GFP was used as a positive control for nuclear localization of the fusion protein (Robatzek and Somssich 2002), whereas a 35S *CaMV*-intron-GFP reporter construct (-GFP) served as a control for non-targeted localization. **(b)** Transient co-transfections of 4-week-old Arabidopsis leaves using the -928 *AtSIRK* promoter-GUS reporter construct (*AtSIRKp::GUS*) together with the various effectors indicated below each cut-out. All effector constructs were under the control of the constitutive 35S *CaMV* promoter. For each construct low magnification leaf sections and higher magnifications of individual bombarded areas is shown. Enlarged cut-out at the bottom right-hand side originates from similar experiments using a 4×W2 synthetic promoter-GUS reporter (Rushton et al. 2002) and the WRKY11-GFP as effector. Bar scale shown in high magnification picture of HA-GFP corresponds to 100 μ m and the same for all pictures of the series

demonstrated nuclear localization of the respective products (Fig. 5a). The same constructs were then used as effectors together with a GUS reporter construct driven by the -928 bp *AtSIRK* promoter (Robatzek and Somssich 2002) for co-bombardment assays using green 4-week-old Arabidopsis leaves. As shown in Fig. 5b, all WRKY fusion proteins, with the exception of WRKY11, were able to activate the reporter gene, demonstrating transactivation capabilities of these transcription factors. WRKY6 (positive control) was previously shown to activate the *AtSIRK* promoter (Robatzek and Somssich 2002). The fact that WRKY11 does not activate the reporter gene despite its

ability to bind to the *AtSIRK* promoter in vitro (Fig. 2) suggested a possible lack of transactivation capabilities. However, in similar experiments in which the GUS reporter gene was driven by the 4×W2 box synthetic promoter (Rushton et al. 2002), the same WRKY11 construct clearly activated the reporter (Fig. 5b, 4×W2::GUS) as did all other tested WRKY constructs (data not shown). Activation by WRKY11 was also observed using the 1×W2m11 sequence (see Fig. 1a) to drive reporter gene expression, but not when the 1×W2m13 sequence or a mutated version of the W2 sequence (4×W2mut::GUS) were used (data not shown). Thus, WRKY11 appears to have inherent transactivation function that requires W-box specific binding in the proper DNA context.

The occurrence of WRKY factor binding sites within the Arabidopsis genome

From our binding studies it is evident that the minimal W-box motif, TTGAC-C/T, is insufficient to define a functional element capable of WRKY factor binding. Based on the information gained in this study we deduced two extended TGAC core-containing weight matrices (Fig. 6a). One, consensus A (TWGTTGACYWWWW), represents a more stringent version of the second, consensus B (DDTTGACYHND). In Arabidopsis, known functional *cis*-regulatory elements appear to be enriched close to gene sequences (Berendzen et al. 2006). Therefore, we used the Motif Mapper Program (<http://www.mpiz-koeln.mpg.de/coupland/coupland/mm3/html/>) to search the Arabidopsis genome for the occurrence of these two consensus sequences. The frequency of these motifs was calculated for three datasets: (a) within the first 600 bp of all annotated gene promoters (relative to the translation start ATG), (b) within the first 1500 bp of all annotated gene promoters and (c) throughout the entire genome sequence. Consensus A was found 174, 393 and 966 times within the respective datasets. This motif was significantly ($P < 0.001$) enriched within the 600 bp promoter versus the whole genome dataset (1.4-fold, Fig. 6b), whereas within the 1500 bp promoter dataset no significant enrichment was observed. With consensus B or the TTGACY hexanucleotide motif, no significant enrichment, compared to the whole genome, was found within the two promoter sets. A functional classification using the Gene Ontology (GO) annotation at TAIR (http://www.arabidopsis.org/portals/genAnnotation/functional_annotation/go.jsp) of the genes containing consensus A in their 600 bp promoter are significantly enriched for proteins that are involved in stress responses or metabolism ($P < 0.001$; Fig. 6c). They are involved in diverse plant functions (Fig. 6c) and include numerous transcription factors and other stress-related proteins (see Supplementary Table 2).

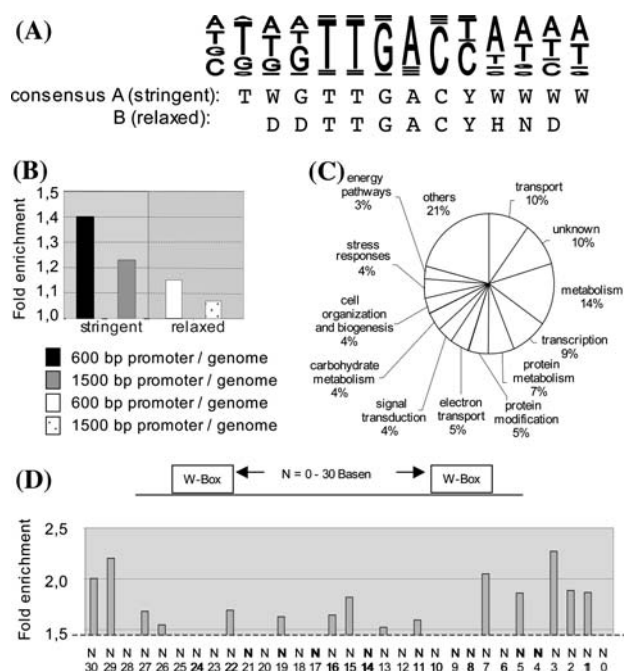


Fig. 6 Motif analysis of WRKY binding sites. **(a)** Sequence logos (Crooks et al. 2004) depicting nucleotide distribution for the WRKY site as derived from our study. The consensus (A) and (B) given below represent a more stringent and a more relaxed version, respectively. **(b)** Fold enrichment of the two consensus sequences within: **(a)** the dataset 600 bp upstream region (relative to the ATG) of all annotated genes versus entire genome sequence, and **(B)** the dataset 1500 bp upstream region (relative to the ATG) of all annotated genes versus entire genome sequence using the program Motif Mapper. **(c)** Pie chart display of the GO functional classification of annotated genes showing an enrichment of WRKY binding sites in their 600 bp upstream sequences identified in **b**. The percentage for each category is shown. **(d)** Fold enrichment of closely spaced WRKY binding sites, ranging from 0 to 30 bp apart, within our 600 bp promoter dataset identified using the Motif Mapper program. The dashed line on the x-axis represents the cut-off value for the statistical background distribution. A 1.4-fold enrichment over the genomic background frequency corresponds to a significant overrepresentation of the motifs ($P < 0.001$). W = A,T; Y = C,T; D = A,G,T; H = A,C,T; N = A,C,T,G

Again using the Motif Mapper program, we investigated whether the appearance of two closely adjacent W-box elements show enrichment in our selected datasets and if so, is there a certain spacing preference between individual motifs. As a framework, we choose to analyze a spacing range from 0 to 30 bp, a potentially important distance inferred from several previous publications (see above). Figure 6d summarizes our findings for the 600bp promoter dataset versus the whole genome. This data implies that certain defined distances are significantly enriched ($P < 0.001$), in particular, closely adjacent (N1–N3, N5, N7) and more distantly spaced (N29–30) W-box elements occur substantially more often than expected.

Discussion

The availability of wholly sequenced genomes in combination with molecular and functional genomic techniques is enabling us to decipher the complete gene-encoded protein set of single organisms. In higher eukaryotes, including plants, 8–10% of such proteins are transcription factors that play a pivotal role in establishing transcriptional regulatory networks governing diverse cellular expression profiles. A key component in understanding such networks is to identify transcription factor binding sites common to the individual genes within defined regulons. Computational methods are currently being applied to assist in locating such sites within entire genomes (Pilpel et al. 2001; Bulyk 2003; Rombauts et al. 2003; Wasserman and Sandelin 2004). However, in order to improve computational predictions of TF sites within promoters and thereby to extract biologically meaningful information, it is absolutely essential to experimentally define more precisely what constitutes a functional *cis*-regulatory element.

With respect to the WRKY transcription factor DNA binding site, our study demonstrates that the previously defined W-box consensus, T/CTGACC/T (Ülker and Somssich 2004), alone is insufficient to predict that it will be bound by these proteins. Rather, as illustrated in Fig. 1, additional neighboring nucleotides also contribute in determining high affinity binding in vitro. Moreover, these nucleotides partly determine the type of WRKY factor that will be recruited. *At*WRKY6 and *At*WRKY11 bind well to W-boxes that have a G residue directly 5' adjacent to the element, whereas *At*WRKY 26, 38 and 43 bind to the same motif if this residue is a T, C or A. This is the first time that such discriminatory WRKY factor binding to W-boxes has been observed. Sequences carrying the hexamer, CTGACC, were not bound by any of the five WRKY representatives tested, indicating that a minimal W-box element should rather be defined as, 5'-TTGACC/T-3'. The sequence motif TTGACA has previously been referred to as a W-box-like motif (Maleck et al. 2000; Kankainen and Holm 2004; Navarro et al. 2004). However, neither our data nor earlier work (de Pater et al. 1996 #3286) lends support to this assumption. In addition, our findings imply that the hexamers TTGACC and TTGACT are not functionally identical with respect to WRKY factor binding. This observation is supported by previous studies demonstrating that the affinity of Arabidopsis ZAP1 (=AtWRKY1) to the TTGACC motif is four-fold higher than to TTGACT (de Pater et al. 1996). By using DNA regions derived from different gene promoters we were able to show that the sequence environment into which W-box elements are embedded can significantly influence protein binding. This influence cannot always be explained merely by

differences in the immediately adjacent neighboring bases, indicating that additional, as of yet ill-defined parameters, also play a role. Nevertheless, based on our results, we derived at two consensus sequences that currently best define good WRKY factor binding sites (Fig. 6).

In several cases closely adjacent W-box elements have been observed in various gene promoters (Eulgem et al. 1999; Yang et al. 1999; Yu et al. 2001; Chen and Chen 2002; Marè et al. 2004; Zhang et al. 2004). In case of *PcWRKY1*, the presence of these multiple sites appears to have a synergistic effect on transcription (Eulgem et al. 1999). On the other hand, the barley *Hv-WRKY38* factor actually requires two closely adjacent W-boxes for efficient DNA binding (Marè et al. 2004). Scanning the entire Arabidopsis genome for two W-box elements from 0 to 30 nucleotides apart revealed that there is a statistically significant enrichment of such dimers within promoter regions. However, although certain distances appear to be favored, the biological importance of this observation remains to be tested. Some WRKY family members do contain leucine zippers capable of forming homo- and hetero-dimers (Cormack et al. 2002; Robatzek and Somssich 2002; Xu et al. 2006; Shen et al. 2007) but the majority do not and very likely bind as monomers to W-box elements. Our data are in agreement with this and indicate that no synergistic binding effects are observed for the tested WRKY factors in the presence of two closely adjacent W-box elements (Fig. 3). These results are further substantiated by recent surface plasmon resonance and NMR studies demonstrating that *AtWRKY4*-CTD binds to the W-box element with a stoichiometry of 1:1 (Yamasaki et al. 2005) and by high-resolution crystal structure of *AtWRKY1*-CTD (Duan et al. 2007). One should note however, that some WRKY factors can influence expression of specific genes without directly contacting DNA. Recently, the rice WRKY factor *OsWRKY51*, although failing to bind itself, was shown to enhance specific binding of *OsWRKY71* to the *Amy32b* gene promoter thereby synergistically suppressing gibberellic acid-inducibility of a tested reporter gene (Xie et al. 2006).

Another important finding is that Arabidopsis members of all three major WRKY groups show W-box specific binding both in vitro and in vivo. No subgroup specific differences were observed although this may have been expected based on the fact that group III WRKY proteins very likely utilize different basic amino acid residues within their DNA binding domains for contacting DNA phosphate groups compared to the other two groups (Yamasaki et al. 2005). These results are in accord with in silico data implying that the lineage-specific expansion of the WRKY domains in Arabidopsis has diversified primarily in terms of their expression patterns rather than in their target DNA-binding sites (Babu et al. 2006). All

tested Arabidopsis WRKY factors localized to the nucleus, which is fully consistent with previous studies on different plant WRKY factors with the exception of *AtWRKY46* whose localization appears to be restricted to the nucleolus (Koroleva et al. 2005).

Using an alanine scanning approach to analyze the C-terminal domain of tobacco *NtWRKY-9*, Maeo et al. (2001) and a site directed mutagenesis approach for the *AtWRKY1*-CTD (Duan et al. 2007) revealed that not all amino acid residues within the highly conserved WRKYGQK peptide stretch of the DNA binding domain are absolutely essential for W-box binding. The substitutions R→A and R→E, Y→F, G→A, Q→A and Q→K still enabled binding to the DNA element in EMSA, albeit at lower efficiencies. One apparent discrepancy between our results and those from Maeo et al. (2001) relate to the substitution of a critical cysteine involved in forming the zinc finger. Whereas this substitution completely abolished W-box binding of *NtWRKY-9* (Maeo et al. 2001), this was not completely the case for *AtWRKY11*-CTD (Fig. 4). One plausible explanation for this difference is that no additional neighboring cysteine or histidine residues are present in the *NtWRKY-9* protein, while two additional histidine residues are present in *AtWRKY11*-CTD, which may partly allow such a finger to form.

Our analysis, that included substitutions at other conserved amino acid positions within the WRKY domain of *AtWRKY11*, suggest that in most cases even significantly conservative exchanges perturbed efficient binding to the W-box (Fig. 4). This hints towards a need for a rather stringent conformational structure for high affinity binding that is also supported by the studies of Duan et al. (2007). Consistent with this assumption is the fact that an Arabidopsis mutant encoding an *AtWRKY52* protein having a single amino acid insertion within the WRKY domain could no longer bind a W-box element (Noutoshi et al. 2005). It is also partly substantiated by protein structure prediction programs and by the NMR solution structure of the C-terminal *AtWRKY4* DNA demonstrating that this region consists of four β -strands forming a compact anti-parallel β -sheet (Yamasaki et al. 2005) with many of the introduced substitutions that negatively affect binding located within the β -sheet strands in WRKY11 CTD (for example m10–m15 and m18 in Fig. 4a). Two of the three substitutions that did not affect WRKY11 CTD binding are outside of the predicted β -strand structural elements (Fig. 4b, m7 and m16). A structural model predicts that the β -strand containing the WRKYGQK motif makes contact with a 6 bp DNA region, consistent with the length of the W-box, in the major DNA groove (Yamasaki et al. 2005; Duan et al. 2007). Verification of this model however, awaits a structural determination of the WRKY protein complexed with its DNA binding site.

The high-resolution crystal structure of AtWRKY1-C revealed an additional β -strand N-terminal to the WRKY domain that was missing in the NMR AtWRKY4-C model (Yamasaki et al. 2005; Duan et al. 2007). Within the loop between this and the adjacent β -strand containing the WRKYGQK motif, a pivotal residue, D308, was identified that forms a well-defined salt bridge with a lysine residue and extensive H-bonds with two additional amino acids. All these residues are conserved among the WRKY proteins and may be important in stabilizing this domain structure (Duan et al. 2007). These stabilizing requirement may explain why the D to E substitution in AtWRKY11 significantly affected binding of the protein to DNA (Fig. 4c, m6).

The co-bombardment assays indicate that all tested WRKY factors has intrinsic transactivation capabilities in vivo (Fig. 5). However, as illustrated by AtWRKY11, W-box binding alone is not sufficient for this function. Very likely promoter architecture as well as additional associated factors will in part determine distinct transcriptional outputs. AtWRKY11 along with its closely related family members AtWRKY7 and AtWRKY17, have been demonstrated to act as negative regulators of defense gene expression in vivo (Journot-Catalino et al. 2006; Kim et al. 2006). Thus, similar to what has been observed for AtWRKY6 (Robatzek and Somssich 2002), these factors may possess dual functionalities dependent on promoter context. No role has as yet been identified for the other here studied factors, AtWRKY26, 38 and 43.

It is evident that our studies cannot fully explain the discrete binding site selectivity of the large set of WRKY factors to their in vivo target sites. Still, as a starting point, our analysis provides important information on what actually constitutes a W-box-like element within regulatory sequences that can be predicted to be bound by certain WRKY family members. These studies therefore should help to improve whole genome in silico analyses regarding WRKY protein-DNA interactions.

Acknowledgements We thank Dr. Aifen Zhou for providing the WRKY6-GFP and WRKY11-GFP expression constructs, and Dr. Janna Brümmer for providing the WRKY26-GFP construct. This work was supported by the DFG-funded Graduiertenkolleg für Molekulare Analysen von Entwicklungsprozessen III/GK-GRK 306/1 and by the DFG/AFGN grants SO 235/3-2/5-4 to Imre E. Somssich.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Asai T, Tena G, Plotnikova J, Willmann MR, Chiu W-L, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J (2002) MAP kinase signalling cascade in *Arabidopsis* innate immunity. *Nature* 415:977–983
- Babu MM, Iyer LM, Balaji S, Aravind L (2006) The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic Acids Res* 34:6505–6520
- Berendzen K, Stuber K, Harter K, Wanke D (2006) *Cis*-motifs upstream of the transcription and translation initiation sites are effectively revealed by their positional disequilibrium in eukaryote genomes using frequency distribution curves. *BMC Bioinformatics* 7:522
- Bulyk ML (2003) Computational prediction of transcription-factor binding site locations. *Genome Biol* 5:201
- Chen C, Chen Z (2000) Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. *Plant Mol Biol* 42:387–396
- Chen C, Chen Z (2002) Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced *Arabidopsis* transcription factor. *Plant Physiol* 129:706–716
- Cormack RS, Eulgem T, Rushton PJ, Köchner P, Hahlbrock K, Somssich IE (2002) Leucine zipper containing WRKY proteins widen the spectrum of immediate early elicitor-induced WRKY transcription factors in parsley. *Biochim Biophys Acta* 1576:92–100
- Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190
- de Pater S, Greco V, Pham K, Memelink J, Kijne J (1996) Characterization of a zinc-dependent transcriptional activator from *Arabidopsis*. *Nucleic Acids Res* 24:4624–4631
- Duan M-R, Nan J, Liang Y-H, Mao P, Lu L, Li L, Wei C, Lai L, Li Y, Su X-D (2007) DNA binding mechanism revealed by high resolution crystal structure of *Arabidopsis thaliana* WRKY1 protein. *Nucleic Acids Res* 35:1145–1154
- Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. *Curr Opin Plant Biol* 10:366–371
- Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE (1999) Early nuclear events in plant defense: rapid gene activation by WRKY transcription factors. *EMBO J* 18:4689–4699
- Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. *Trends Plant Sci* 5:199–206
- Gonzalez-Lamothe R, Tsitsigiannis DI, Ludwig AA, Panicot M, Shirasu K, Jones JDG (2006) The U-Box protein CMPG1 is required for efficient activation of defense mechanisms triggered by multiple resistance genes in tobacco and tomato. *Plant Cell* 18:1067–1083
- Hamberger B, Hahlbrock K (2004) The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc Natl Acad Sci USA* 101:2209–2214
- Hara K, Yagi M, Kusano T, Sano H (2000) Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor on wounding. *Mol Gen Genet* 263:30–37
- Heise A, Lippok B, Kirsch C, Hahlbrock K (2002) Two immediate-early pathogen-responsive members of the AtCMPG gene family in *Arabidopsis thaliana* and the W-box-containing elicitor-response element of AtCMPG1. *Proc Natl Acad Sci USA* 99:9049–9054
- Hinderhofer K, Zentgraf U (2001) Identification of a transcription factor specifically expressed at the onset of leaf senescence. *Planta* 213:469–473
- Johnson CS, Kolevski B, Smyth DR (2002) TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *Plant Cell* 14:1359–1375

- Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in *Arabidopsis thaliana*. *Plant Cell* 18:3289–3302
- Kankainen M, Holm L (2004) POBO, transcription factor binding site verification with bootstrapping. *Nucleic Acids Res* 32:222–229
- Ke S-H, Madison EL (1997) Rapid and efficient site-directed mutagenesis by single-tube ‘megaprimer’ PCR method. *Nucleic Acids Res* 25:3371–3372
- Kim K-C, Fan B, Chen Z (2006) Pathogen-induced *Arabidopsis* WRKY7 is a transcriptional repressor and enhances plant susceptibility to *Pseudomonas syringae*. *Plant Physiol* 142:1180–1192
- Koroleva OA, Tomlinson ML, Leader D, Shaw P, Doonan JH (2005) High-throughput protein localization in *Arabidopsis* using *Agrobacterium*-mediated transient expression of GFP-ORF fusions. *Plant J* 41:162–174
- Li J, Brader G, Kariola T, Tapio Palva E (2006) WRKY70 modulates the selection of signaling pathways in plant defense. *Plant J* 46:477–491
- Luo M, Dennis ES, Berger F, Peacock WJ, Chaudhury A (2005) MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in *Arabidopsis*. *Proc Natl Acad Sci USA* 102:17531–17536
- Maeo K, Hayashi S, Kojima-Suzuki H, Morikami A, Nakamura D (2001) Role of conserved residues of the WRKY domain in the DNA-binding of tobacco WRKY family proteins. *Biosci Biotechnol Biochem* 65:2428–2436
- Maleck K, Levine A, Eulgem T, Morgen A, Schmid J, Lawton K, Dangel JL, Dietrich RA (2000) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nature Genet* 26:403–410
- Marè C, Mazzucotelli E, Crosatti C, Francia E, Stanca MA, Cattivelli L (2004) Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. *Plant Mol Biol* 55:399–416
- Navarro L, Zipfel C, Rowland O, Keller I, Robatzek S, Boller T, Jones JDG (2004) The transcriptional innate immune response to flg22. Interplay and overlap with avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol* 135:1113–1128
- Noutoshi Y, Ito T, Seki M, Nakashita H, Yoshida S, Marco Y, Shirasu K, Shinozaki K (2005) A single amino acid insertion in the WRKY domain of the *Arabidopsis* TIR-NBS-LRR-WRKY-type disease resistance protein SLH1 (sensitive to low humidity 1) causes activation of defense responses and hypersensitive cell death. *Plant J* 43:873–888
- Pilpel Y, Sudarsanam P, Church GM (2001) Identifying regulatory networks by combinatorial analysis of promoter elements. *Nat Genet* 29:153–159
- Robatzek S, Somssich IE (2002) Targets of AtWRKY6 regulation during plant senescence and pathogen defense. *Genes Dev* 16:1139–1149
- Rombauts S, Florquin K, Lescot M, Marchal K, Rouzé P, Van de Peer Y (2003) Computational approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant Physiol* 132:1162–1176
- Ross CA, Liu Y, Shen QJ (2007) The WRKY gene family in rice (*Oryza sativa*). *J Int Plant Biol* 49:827–842
- Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE (1996) Interaction of elicitor-induced DNA binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J* 15:5690–5700
- Rushton PJ, Reinstädler A, Lipka V, Lippok B, Somssich IE (2002) Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen- and wound-induced signaling. *Plant Cell* 14:749–762
- Shen Q-H, Saijo Y, Mauch S, Biskup C, Bieri S, Keller B, Seki H, Ulker B, Somssich IE, Schulze-Lefert P (2007) Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* 315:1098–1103
- Sun C, Palmqvist S, Olsson H, Borén M, Ahlandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the *iso1* promoter. *Plant Cell* 15:2076–2092
- Ülker B, Somssich IE (2004) WRKY transcription factors: from DNA binding towards biological function. *Curr Opin Plant Biol* 7:491–498
- van Verk MC, Pappaioannou D, Neeleman L, Bol JF, Linthorst HJM (2008) A novel WRKY transcription factor is required for induction of *PR-1A* gene expression by salicylic acid and bacterial elicitors. *Plant Physiol* 146:1983–1995
- Wang D, Amornsiripanitch N, Dong X (2006) A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS Pathog* 2:e123
- Wang Z, Yang P, Fan B, Chen Z (1998) An oligo selection procedure for identification of sequence-specific DNA-binding activities associated with plant defense. *Plant J* 16:515–522
- Wasserman WW, Sandelin A (2004) Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* 5:276–287
- Xie Z, Zhang Z-L, Zou X, Yang G, Komatsu S, Shen QJ (2006) Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells. *Plant J* 46:231–242
- Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. *Plant Cell* 18:1310–1326
- Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Tomo Y, Hayami N, Terada T, Shirouzu M, Tanaka A, Seki M, Shinozaki K, Yokoyama S (2005) Solution structure of an *Arabidopsis* WRKY DNA binding domain. *Plant Cell* 17:944–956
- Yang P, Wang Z, Fan B, Chen C, Chen Z (1999) A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of tobacco class I chitinase gene promoter. *Plant J* 18:141–149
- Yu D, Chen C, Chen Z (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of *NPR1* gene expression. *Plant Cell* 13:1527–1539
- Zhang Z-L, Xie Z, Zou X, Casaretto J, Ho T-hD, Shen QJ (2004) A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiol* 134:1500–1513
- Zheng Z, Qamar SA, Chen Z, Mengiste T (2006) *Arabidopsis* WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48:592–605